Online Gradient Descent

(Online Perceptron Algorithms)

$$f^{(t)}(\boldsymbol{w}) = \boldsymbol{w} \cdot \boldsymbol{x}^{(t)}$$

linear predictor (loss)

$$(\boldsymbol{w}) = \frac{1}{\eta}||\boldsymbol{w}||_2^2$$

Euclidean regularization

$$\boldsymbol{w}^{(t+1)} = \boldsymbol{w}^{(t)} - \eta \nabla \ell(y^{(t)}, \hat{y}^{(t)})$$

Additive update

$$BL\sqrt{2T}$$

No Regret

Online Exponentiated Gradient Descent

(Hedge/GWM, Winnow)

$$f^{(t)}(\boldsymbol{w}) = \boldsymbol{w} \cdot \boldsymbol{x}^{(t)}$$

linear predictor (loss)

$$(\boldsymbol{w}) = \frac{1}{\eta}\sum_n w_n \log(w_n)$$

Entropic Regularization

$$\boldsymbol{w}^{(t+1)} = \boldsymbol{w}^{(t)} \exp\left(-\eta \nabla \ell(y^{(t)}, \hat{y}^{(t)})\right)$$

Exponential update

$$\sqrt{(T/2)\log N}$$

No Regret